



Published in final edited form as:

*J Am Acad Child Adolesc Psychiatry*. 2015 July ; 54(7): 541–549. doi:10.1016/j.jaac.2015.04.007.

## Randomized Controlled Trial of Therapeutic Horseback Riding in Children and Adolescents With Autism Spectrum Disorder

Dr. Robin L. Gabriels, PsyD, Dr. Zhaoxing Pan, PhD, Ms. Briar Dechant, BS, Dr. John A. Agnew, PhD, Ms. Natalie Brim, BA, and Dr. Gary Mesibov, PhD

Drs. Gabriels, Pan, and Ms. Dechant are with the University of Colorado Anschutz Medical Campus and the Children's Hospital Colorado, Aurora, CO. Dr. Agnew and Ms. Brim completed this work while with the University of Colorado Anschutz Medical Campus and the Children's Hospital Colorado. Dr. Mesibov is with the University of North Carolina, Chapel Hill

### Abstract

**Objective**—This study expands previous equine-assisted intervention research by evaluating the effectiveness of therapeutic horseback riding (THR) on self-regulation, socialization, communication, adaptive, and motor behaviors in children with autism spectrum disorder (ASD).

**Method**—Participants with ASD (ages 6–16 years; N=127) were stratified by nonverbal IQ standard scores ( $< 85$  or  $> 85$ ) and randomized to one of two groups for 10 weeks: THR intervention or a barn activity (BA) control group without horses that employed similar methods. The fidelity of the THR intervention was monitored. Participants were evaluated within one month pre- and post-intervention by raters blind to intervention conditions and unblinded caregiver questionnaires. During the intervention, caregivers rated participants' behaviors weekly.

**Results**—Intent-to-treat analysis conducted on the 116 participants who completed a baseline assessment (THR n = 58; BA control n = 58) revealed significant improvements in the THR group compared to the control on measures of irritability (primary outcome) ( $p=.002$ ; effect size [ES]=.50) and hyperactivity ( $p=.001$ ; ES=0.53), beginning by week five of the intervention. Significant improvements in the THR group were also observed on a measure of social cognition ( $p=.05$ , ES=.41) and social communication ( $p=.003$ ; ES=.63), along with the total number of words ( $p=.01$ ; ES=.54) and new words ( $p=.01$ ; ES=.54) spoken during a standardized language sample. Sensitivity analyses adjusting for age, IQ, and per-protocol analyses produced consistent results.

© 2015 American Academy of Child and Adolescent Psychiatry.

Correspondence to Robin L. Gabriels, PsyD, Department of Psychiatry and Behavioral Sciences, University of Colorado Anschutz Medical Campus/Children's Hospital Colorado, 13123 E. 16th Ave, B130, Aurora, CO 80045; robin.gabriels@childrenscolorado.org.

Disclosure: Dr. Gabriels is a co-author of the book *Growing Up With Autism: Working With School-Age Children and Adolescents* (Guilford Press), and the book *Autism - From Research to Individualized Practice* (Jessica Kingsley Publishers), from which she has received royalties. She has received grant funding from the Simons Foundation, Nancy Lurie Marks Family Foundation, MARS/WALTHAM, and the Human-Animal Bond Research Institute (HABRI) Foundation. She has served on the Medical Advisory Board of Professional Association of Therapeutic Horsemanship International and has received honoraria from Spring Harbor Hospital, Portland, ME, Indiana University Neuroscience Center, Indianapolis, IN, and Routledge/Taylor and Francis Group. Dr. Mesibov is the co-author of the book *The TEACCH Approach to Autism Spectrum Disorders* (Springer Press), from which he has received royalties. Drs. Pan and Agnew, Ms. Dechant, and Ms. Brim report no biomedical financial interests or potential conflicts of interest.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Conclusion**—This is the first large-scale randomized, controlled trial demonstrating efficacy of THR for the ASD population, and findings are consistent with previous equine-assisted intervention studies.

**Clinical trial registration information**—Trial of Therapeutic Horseback Riding in Children and Adolescents With Autism Spectrum Disorder; <http://clinicaltrials.gov/>; NCT02301195.

### Keywords

Autism spectrum disorder; equine-assisted activities and therapies; human-animal interaction; therapeutic horseback riding; social-communication functioning

## INTRODUCTION

Autism spectrum disorder (ASD) is a neurodevelopmental disorder diagnosed based on core impairments in social interaction, communication abilities, and the presence of restricted, repetitive, and stereotyped behaviors and interests.<sup>1</sup> Although these core impairments have a heterogeneous clinical presentation across individuals with ASD, limited joint attention skills (i.e., being aware of and sharing another's focus of attention) appear to be an important issue underlying social-communication symptoms.<sup>2</sup> Improving joint attention skills in the ASD population has been linked to better behavioral regulation, communication skills, and social competence.<sup>3</sup>

Children with ASD exhibit more challenging or inappropriate behaviors than typically developing peers or those diagnosed with other psychopathology.<sup>4–6</sup> The ASD population has high incidences of stress, anxiety, and depression,<sup>7,8</sup> along with a variety of other challenges such as irritability, hyperactivity, and problems with sensory processing, dyspraxia, intellectual functioning, adaptive daily living skills, and sleep.<sup>9–11</sup> Co-occurrence of these issues with core ASD symptoms can present challenging behaviors, often persisting beyond childhood into adolescence and adulthood.<sup>12</sup> Research indicates that care and management of these challenging behaviors contributes to higher rates of stress and greater burden among caregivers of individuals with ASD compared with other special needs populations.<sup>13–15</sup>

Incorporating animals into the treatment process to decrease problem behaviors and improve functioning has been proposed as a promising area of intervention for these at-risk populations.<sup>16</sup> Evidence for animal-assisted interventions (AAI) is limited; however, including animals in health care settings has been hypothesized to produce sensory-related relaxation experiences that allow children to better manage stressful events and engage in prosocial behaviors<sup>17</sup>. Studies have cited reductions in stress hormone levels (i.e. cortisol) in children with ASD following interactions with service dogs, and also in a general child population following an 11-week equine intervention.<sup>18, 19</sup> The ASD population has a particular need for interventions that target reductions in stress-related maladaptive behaviors and improvements in social-communication functioning.

O'Haire's (2013) comprehensive review of AAIs targeting individuals with ASD identified only 14 empirically-based studies from 1840 to 2011, with interventions that intentionally

included a live animal, and reported results. In those studies, results indicated improvements in social interaction communication skills, behaviors, and stress levels, even though intervention methods and types of animals varied (i.e., dogs, guinea pigs, lamas, rabbits, horses). O’Haire (2013) recommended future AAI studies increase methodological rigor by randomizing participants to comparison treatments without animals, employing outcome raters blind to intervention conditions, and replicating methods.<sup>20</sup>

One form of AAI includes Equine-Assisted Activities and Therapies (EAAT). This is a rapidly growing field for special needs populations that involves horses in a therapeutic intervention setting. Since 2009, there has been an increase in EAAT publications citing various methods for working with the ASD population, including therapeutic horseback riding (THR),<sup>21–24</sup> hippotherapy (HPOT),<sup>25,26</sup> psychoeducational horseback riding (PER),<sup>27</sup> and EAAT.<sup>28, 29</sup> Although each method offers a different therapeutic focus and type of activity, the inclusion of horses as part of the therapeutic experience is a common thread. THR typically involves riding in small group settings led by a certified THR instructor teaching horsemanship skills targeting therapeutic goals.<sup>30</sup> Hippotherapy is led by an occupational, speech, or physical therapist, and intervention activities use the movement of the horse to target functional outcomes.<sup>31</sup>

Though EAAT methods vary, cited improvements with the ASD population overlap across studies, suggesting that some aspect(s) of the human-equine interaction may be important for change. Temple Grandin, an accomplished adult with ASD, asserts that one key element of horseback riding that helped her decrease anxiety-related feelings and behaviors involves the reciprocal/joint attention relationship, or “teamwork,” between the rider and the horse (e.g., moving “in sync with the horse’s body” and the horse’s intuitive sensitivity and responsiveness to the rider) (p. 6).<sup>32</sup> Despite EAAT study variations (i.e., duration, activities, and targeted outcome measures), results include decreases in problem behaviors (e.g., irritability, hyperactivity, stereotypy, inattention, self-injury, and ASD symptom severity)<sup>21–23,26,28</sup> as well as increases in sensory processing,<sup>21, 28, 29</sup> communication,<sup>22, 23, 25, 27, 29</sup> and social skills.<sup>21, 27–29</sup> Although these EAAT studies offer promising preliminary effects for individuals with ASD, generalizability of these findings are limited by small sample sizes, minimal methods to confirm the ASD diagnosis, and use of potentially biased informants for outcome measures.

The present study builds upon the piloted<sup>22</sup> 10-week THR intervention curriculum conducted by the authors and available from the first author. The objective of this RCT was to evaluate whether the THR could affect significant improvements on measures of self-regulation, communication, social, adaptive, and motor behaviors in children and adolescents diagnosed with ASD. The second, exploratory, aim to evaluate the retention of improvements six months following the THR intervention phase will be reported separately. This study has an elevated level of methodological rigor (i.e., randomized controlled design, use of a well-established standardized ASD diagnostic measure, control group without horses that mirrored THR intervention methods, fidelity measure of the THR intervention, and outcome evaluators blind to participants’ intervention condition), and was conducted at an established (30 years) riding center site with “premiere” certification by PATH

International,<sup>30</sup> meaning the center adheres to the highest safety and ethical standards in the industry for all living beings involved in the treatment process.

## METHOD

### Participants

This study was approved by the Institutional Review Board (IRB) at the first author's institution. Participants were recruited using IRB-approved fliers distributed to university-affiliated hospitals, schools, ASD parent support organizations, and community providers. Participants' caregivers were compensated for mileage to study visits.

Participants were included in the study if they met the following criteria: aged 6 to 16 years; met or exceeded the ASD screening cut-off ( $\geq 15$ ) on the Social Communication Questionnaire (SCQ);<sup>33</sup> had an ASD diagnosis, confirmed by meeting clinical cut-offs for ASD on the Autism Diagnostic Observation Schedule (ADOS) or ADOS-second edition (ADOS-2);<sup>34, 35</sup> had a combined score on the Irritability and Stereotypy subscales of the Aberrant Behavior Checklist-Community (ABC-C)<sup>36</sup> of  $\leq 11$  and had a Leiter-R<sup>37</sup> Brief nonverbal IQ (NVIQ) standard score of  $\geq 40$ , as established by the pilot.<sup>22</sup> Exclusion criteria included: a previously identified genetic disorder known to be causative of or resulting in a phenotype similar to ASD, a history of medical or behavioral issues making participation dangerous, a history of animal abuse or phobia of horses, more than two hours of EAAT within the past six months, or weight exceeding the riding center's policies to ensure the health and safety of staff and volunteers. In cases with more than one child with ASD in a family, only the first sibling who qualified for the study was included to avoid duplication of caregiver reporting styles. See Figure 1 for the participant flowchart.

### Study Design

**Screening**—Participants were scheduled for an initial study diagnostic and NVIQ assessment visit at the first author's institution, during which informed consent from each participant's legal guardian was obtained. If the participant was seven or older, verbal, and had an IQ  $\geq 80$ , informed assent was obtained (and child's wishes were not overridden).

**Randomization**—Eligible participants were randomized into one of two groups (THR or BA). Random group assignment study numbers were generated by the statistician on this project using a size-four block randomization stratified by NVIQ ( $\leq 85$  or  $> 85$ ) to ensure that this factor was comparable between groups. Participants randomized to the barn activity (BA) control group were offered two free riding lessons at study conclusion.

**Riding Center Screening**—After physician and caregiver medical clearance forms were completed, participants were screened at the riding center by their assigned group leader. Study attendance and riding center policies were reviewed, and physical, behavioral, and adaptive functioning levels were observed in order to assign groups similar in age and/or ability. THR group participants were fitted for a helmet and mounted a horse for 10 minutes while they engaged in a consistent sequence of activities (i.e., hold reins, asking the horse to “walk on” and “whoa,” stretching exercises, brief seated trot) with the assistance of one

volunteer who led the horse and up to two volunteers who walked beside the horse. BA control group participants completed a brief art project.

### Interventions

Both 10-week interventions were a minimum of 45 minutes long each session, had two to four participants, had equine-related information content, assigned at least one volunteer to each participant, and used behavioral teaching methods commonly used for the ASD population (e.g., use of visual aids, directly praising appropriate behaviors, and using the participant's interest to engage).<sup>38</sup> Every effort was made to maintain the consistency of horses for the THR group and volunteers for the THR and BA control groups.

**Therapeutic Horseback Riding Intervention**—A certified PATH International<sup>30</sup> advanced therapeutic riding instructor taught all lessons and followed the study manual approach that had a two-part teaching focus: 1) therapeutic riding skills (e.g., mounting, halting, steering, turning, and trotting) and 2) horsemanship skills (e.g., how to lead and care for their horse). Lessons followed a consistent routine presented as a picture schedule: 1) put on riding helmet, 2) wait on the bench, 3) mount horse, 4) riding activities, 5) dismount horse, 6) groom horse, and 7) put away equipment. The riding portion consisted of a warm up activity, skill review, learning a new skill, lesson review, and a cool down activity. After riding, participants led their horses to the tacking area where they learned and practiced un-tacking and grooming skills, and thanked their horses and volunteers.

**Barn Activity Control Intervention**—The BA control groups were co-led by a THR instructor and a master's level therapist who had expertise working with and modifying curriculums for children with ASD. Participants had no contact with horses; however, a life-sized stuffed horse was an integral part of teaching horsemanship skills.

**Fidelity**—The essential components of the THR intervention groups were monitored by the principal investigator (PI) or co-investigator (Co-I) scoring 20% of the THR 10-week sessions on a four-point rating scale after they achieved at least 80% inter-reliability. The fidelity instrument targeted eight core areas covering environmental, volunteer, and instructor factors consistent with ASD learning needs.

### Behavioral Outcome Measures

Within one month pre-and post-intervention, a battery of assessment measures were conducted to evaluate participants' baseline and post-intervention functioning levels. A consistent caregiver not blinded to the intervention assignment for each participant completed all behavior rating forms before, during, and after the intervention phase. Study assessment personnel were blind to participants' intervention assignment and did not have access to participants' pre-intervention evaluations when conducting post- evaluations.

A speech therapist measured participants' receptive vocabulary using the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4), norm-referenced and standardized for individuals ages 2 years, 6 months through age 90 years.<sup>39</sup> The PPVT-4 has two parallel forms: form A was used for pre-evaluations, and form B was used for post-evaluations. A

five-minute expressive language sample was elicited from participants using the *Systematic Analysis of Language Transcripts (SALT)*<sup>40</sup>, which provides guidelines for eliciting, transcribing, and analyzing language samples from individuals, including those with ASD. An occupational therapist administered the Bruininks-Oseretsky Test of Motor Proficiency–2<sup>nd</sup> Ed (BOT-2)<sup>41</sup> (Short Form) and two subscales of the Sensory Integration and Praxis Test (SIPT)<sup>42</sup>: Praxis on Verbal Command and Postural Praxis. Participants' adaptive behaviors were evaluated by a study research assistant using the Vineland Adaptive Behavioral Scales – 2<sup>nd</sup> Edition, Survey Interview Form (VABS-II)<sup>43</sup> with caregivers.

Caregivers rated participants' irritability, lethargy/social withdrawal, stereotypy, hyperactivity, and inappropriate speech behaviors via the ABC-C<sup>36</sup>. The ABC-C is a 58-item symptom checklist for assessing problem behaviors of children and adults with developmental disabilities in community settings.<sup>44, 45</sup> Extensive psychometric assessment of the ABC-C indicates that its subscales have high internal consistency, adequate reliability, and established validity.<sup>44, 45</sup> In addition, the ABC-C Irritability subscale is an established outcome measure in randomized clinical trials with the ASD population.<sup>46</sup> Caregivers also completed the Social Responsiveness Scale (SRS), which is a 65-item questionnaire that measures the social impairments in ASD on five subscales (Social Awareness, Social Cognition, Social Motivation, Social Communication and Autistic Mannerisms).<sup>47</sup> The SRS has high internal consistency and retest temporal stability in males and females<sup>48</sup> and was used in a previous THR study.<sup>21</sup>

### Statistical Analyses

SAS 9.3 was used for all the analyses.<sup>49</sup> Student t-tests and Chi square tests were used to compare the participants in the BA control group with the participants in THR intervention group for continuous and categorical demographic along with clinical and baseline outcome variables. The primary analyses used a linear mixed effects model (LMM) consisting of the baseline value and the post-evaluations as outcome measures, evaluation time (baseline or post-evaluation) of outcome, group (THR or BA), and their interaction term as fixed effects and an unstructured covariance. Estimate of the interaction was used to assess the efficacy and test for statistical significance. The primary analysis included any randomized participants (n=116) who had either baseline and/or post-evaluation measure to follow the intent-to-treat (ITT) principle. No data imputation was taken. Sensitivity analysis was used in various ways to reanalyze the outcome found significant from the primary analyses to assess how robust the conclusions were. First, the efficacy was further tested while introducing age, IQ, and gender into the primary model as covariates even if there was no difference between the two groups in these covariates. Second, per-protocol analyses (including participants completing 80% or more of either intervention, n=100) were performed using the same method as the primary analysis. Finally a similar mixed effects model was used to fit the weekly ABC-C data to delineate the time course under which contrasts were used to test the between-group difference in the change from baseline by each week. The potential moderator effect of age, IQ, or gender was respectively examined by the same mixed effects model, with this factor and the three level interaction term of this factor, time, and group as two additional fixed effects. The fidelity of the THR treatment implementation was computed as a percentage of the eight intervention component ratings.

Irritability subscale of ABC-C was deemed as the primary outcome. No adjustment for multiple secondary outcome variables was applied.

### Power of the Study

This study was primarily powered for efficacy tests to examine whether the THR intervention led to greater improvements on the ABC-C irritability subscale as compared to the BA control after the 10-week intervention. Taking into account active control group, the assumed effect size for this study is smaller than that observed in the pilot.<sup>22</sup> A sample size of 116 with equal group split ensures 80% power at 5% significance to detect the significant efficacy of THR versus the BA control using a linear mixed effects model when the effect size is 0.53 or greater of the standard deviation of the post-intervention change from baseline.

## RESULTS

### Preliminary Analyses

Of the 144 potential participants screened, 127 (88%) met study inclusion criteria and were enrolled in the trial and randomized (see Figure 1). The two randomized groups did not differ at baseline (see Table 1). After the intervention was initiated, four participants dropped from the THR group and seven participants dropped from the BA control group. There were no significant demographic differences between the two groups for dropped participants. One THR and two BA participants reported initiation or discontinuation of medications during the intervention phase.

**THR Intervention Fidelity**—The average fidelity rating for the eight intervention components were as follows: volunteers were quiet and attentive to instructor and rider (98.2%); lesson schedule followed (100.0%); picture schedule used (90.6%); minimal environmental disruptions (90.6%); instructors provided specific praise to participants (98.9%); instructors maintained calm demeanor (100.0%); visual cues used for directions (100.0%); volunteers were informed of lesson focus and rider needs (98.9%). Eighty-three percent of THR participants rode the same horse for the full 10-week session.

### Clinical Outcomes

Table 2 shows the efficacy of the THR intervention compared to the BA control group for the primary and secondary outcome variables.

**Self-Regulation (ABC-C)<sup>36</sup>**—Participants in the THR group had significantly more improvements from baseline to post-intervention on the ABC-C Irritability subscale score with an effect size of 0.50 ( $p=.02$ ). The analysis of the ABC-C weekly data indicated that the two groups were significantly different in the change from baseline beginning by the fifth week of the intervention (Figure 2).

Similarly, participants in the THR group showed significantly more improvements in the ABC-C Hyperactivity subscale score with an effect size of 0.53 ( $p=.01$ ). The time course of the weekly profile of Hyperactivity subscale had a similar pattern as the Irritability subscale

with significantly greater improvement in THR group starting at week 5. There was no significant difference between the two groups on the other ABC-C subscales.

**Social Measure (SRS)<sup>47</sup>**—The THR group had a greater improvement on the Social Cognition and Communication subscales of the SRS as compared to the BA control group with effect sizes of 0.41 ( $p=.05$ ) and 0.63 ( $p=.003$ ). There was no significant difference between groups on the other subscales of the SRS.

**Communication (SALT)<sup>40</sup>**—Although participants in the THR group spoke fewer words at the pre-intervention language assessment, this was not significantly different from the BA control group. However, following the intervention phase, the THR group had a significant increase in the use of different words with an effect size of 0.54 ( $p=.01$ ) and spoke more words with an effect size of 0.54 ( $p=.01$ ) post-intervention compared to the BA control group.

**Other Secondary Outcomes**—No statistically significant between-group differences were found with respect to the rest of secondary outcomes, including adaptive (VABS-II)<sup>43</sup> and motor (BOT-2 and SIPT) behaviors,<sup>41,42</sup> and PPVT-4.<sup>39</sup>

**Sensitivity and Moderator Analyses**—After adjusting for age, IQ, and gender in the ITT analysis, significantly greater improvements in THR group remained for Irritability (ES=0.51,  $p=.02$ ), Hyperactivity (ES=0.53,  $p=.02$ ), Social Communication (ES=0.72,  $p=.002$ ), number of words spoken, and number of new words (ES=0.52,  $p=.02$ ), while marginally significant improvement was observed for Social Cognition score (ES=0.52,  $p=.06$ ). In per-protocol (PP) analysis of participants completing 80% of the intervention ( $n=52$  THR,  $n=48$  BA control), greater improvement in the THR group was also observed for Irritability (ES=0.42,  $p=.05$ ), Hyperactivity (ES=0.45,  $p=.04$ ), Social Cognition (ES=0.33,  $p=.11$ ), Social Communication (ES=0.57,  $p=.08$ ), number of words spoken (ES=0.55,  $p=.009$ ), and number of new words (ES=0.56,  $p=.009$ ), leading to similar conclusions as the primary ITT analysis. No significant moderation effect was found for age, IQ or gender. Of note, co-morbid diagnoses at study entry were obtained by caregiver report and not independently confirmed; therefore, these data were not included in the moderator analyses.

## DISCUSSION

This article reports results from a randomized control study of 116 participants diagnosed with ASD (ages 6–16 years) involved in a 10-week THR intervention compared to an activity control group that had no interaction with horses. Results show significant post-intervention improvements in the THR group compared to the BA control on the Irritability and Hyperactivity subscales of the ABC-C<sup>36</sup> beginning by the fifth week of intervention. This finding replicates our pilot study.<sup>22</sup> Compared to the BA control, the THR group also showed significant improvements on the SRS<sup>47</sup> Social Cognition and Communication subscales along with significant improvements in the amount of words and different words spoken during a standard language. The outcomes of this study lend support to findings from previous EAAT studies with the ASD population, suggesting that there is an important



active ingredient in the human-equine interaction that can affect positive changes in irritability, hyperactivity, social, and communication behaviors in this population.

Results generate hypotheses regarding the role of the human-equine interaction requiring further investigation. One hypothesis is that riding and working together with the horse to engage in therapeutic riding activities involves a nonverbal joint-attention or shared attention experience that may serve a platform for improving behaviors and social-communication skills in children with ASD. This nonverbal communication between the horse and the rider may include the fact that horses constantly mirror and respond to the rider's body language.<sup>50</sup> Also, this shared attention experience may be enhanced by the enormity of the horse combined with the task demand for the rider of maintaining bilateral control and balance. Outcome measures of joint attention skills, including semi-structured play assessments and behavioral observation coding, may be useful considerations for future THR studies. A second hypothesis is that the human-equine experience (i.e., warmth of the horse's body and rhythmic movement of riding the horse) promotes a relaxing context, which may have a calming effect on children with ASD. The impact of THR on reducing stress levels can be explored using objective behavioral observation measures combined with biological measures such as galvanic skin response or salivary cortisol.<sup>50</sup> Physiological measures may provide more accurate assessments of the stress levels in the ASD population, as gathering self-report accounts are challenging in this population.

The study is limited by the fact that several outcomes (i.e., ABC-C, SRS, VABS-II) were measured using non-blinded caregiver report measures, which is a unique confound to this type of intervention research and raises questions as to whether knowledge produced a placebo effect. An additional limitation is the lack of objective observational measures, which are needed to gain further insight into the mechanisms of behavioral changes observed in this study. This study also used broad measures of motor coordination, which limited the ability to detect possible motor-related changes. In order to have a comprehensive assessment of the THR intervention, a study with a placebo group (i.e. no intervention at all) in addition to the BA control group will provide a better insight into the effect of THR. Additional future study considerations include expanding demographic measures to include, for example, a broader measure of intellectual functioning. Finally, no adjustment was made for multiple secondary outcome comparisons in the analysis, as this may have increased the type I error rate. Therefore, caution needs to be taken into account when interpreting the secondary outcome results.

This study further establishes the evidence base supporting EAAT as a viable therapeutic option for children and adolescents with ASD. Further research is warranted to examine if the joint attention and movement experiences are key THR mechanisms to observe behavioral and social-communication improvements in the ASD population.

## Acknowledgments

This study was supported by grant R01NR012736 from the National Institute of Nursing Research (R.L.G.). Dr. Pan served as the statistical expert for this research.

The authors gratefully acknowledge the children and families who participated in this study, those who assisted with this project from Children's Hospital Colorado (Laurie Burnside, MSM, CCRC, Tina Farrell, CCC-SLP, Syd

Martin, MS, OTR, Shana Holderness, MA, Oren Gordon, BA, and Jessie Lucas, student research assistant), and personnel from the Colorado Therapeutic Riding Center (Amy Shoffner, MSW, PATH International Advanced instructor, Jody Howard, PATH International Advanced Instructor and Registered Evaluator, and Penelope Powell, BA, PATH International Advanced Instructor and Registered Evaluator). The authors also thank PATH International and its former CEO, Kay Green, BA, for disseminating information about this study along with Luitpold Pharmaceuticals and Allen Mann, Director, Luitpold Animal Health, for donating Adequan for study horses. The authors are grateful to Carol Beresford, MD, and Doug Novins, MD, University of Colorado Anschutz Medical Campus for assistance with manuscript preparation.

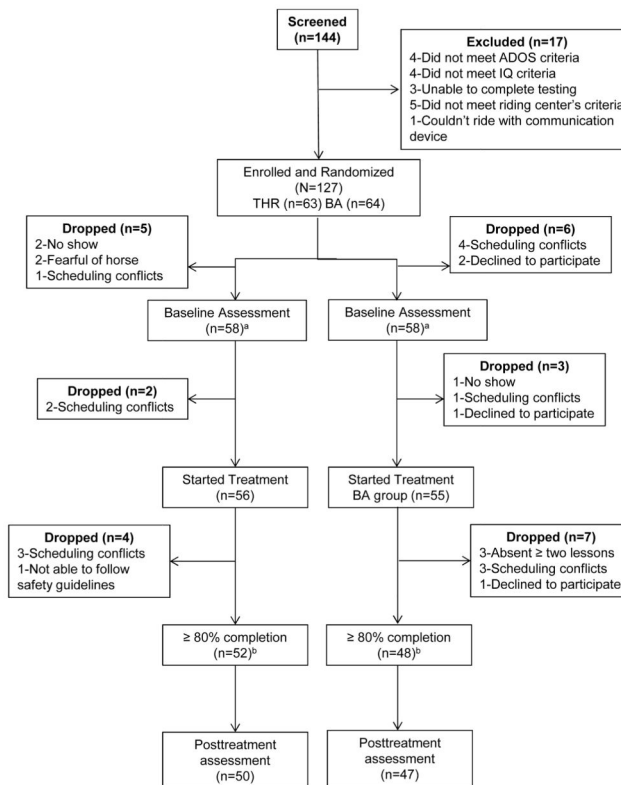
## References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Text Revision (DSM IV – TR). Vol. IV. Washington, D.C: American Psychiatric Association; 2000.
2. Bruinsma Y, Koegel RL, Koegel LK. Joint attention and children with autism: A review of the literature. *Ment Retard Dev Disab Res Rev*. 2004; 10:169–175.
3. Warreyn P, Van Der Paelt E, Roeyers H. Social-communicative abilities as treatment goals for preschool children with autism spectrum disorder: the importance of imitation, joint attention, and play. *Developmental Medicine and Child Neurology*. 2014;1–5.
4. Matson JL, Gonzalez ML, Rivet TT. Reliability of the Autism Spectrum Disorder-Behavior Problems for Children (ASD-BPC). *Research in Autism Spectrum Disorders*. 2008; 2:696–706.
5. Matson JL, Nebel-Schwalm M. Assessing challenging behaviors in children with autism spectrum disorders: a review. *Res Dev Disabil*. 2007; 28(6):567–579. [PubMed: 16973329]
6. Matson JL, Wilkens J, Macken J. The relationship of challenging behaviors to severity and symptoms of Autism Spectrum Disorders. *Journal of Mental Health Research in Intellectual Disabilities*. 2009; 2:29–44.
7. Leyfer OT, Folstein SE, Bacalman S, et al. Comorbid psychiatric disorders in children with autism: interview development and rates of disorders. *J Autism Dev Disord*. 2006; 36:849–61. [PubMed: 16845581]
8. White SW, Oswald D, Ollendick T, Scahill L. Anxiety in children and adolescents with autism spectrum disorders. *Clinical Psychology Review*. 2009; 29(3):216–229. [PubMed: 19223098]
9. Gabriels RL, Cuccaro ML, Hill DE, Ivers BJ, Goldson E. Repetitive behaviors in autism: Relationships with associated clinical features. *Res Dev Disabil*. 2005; 26(2):169–181. [PubMed: 15590247]
10. Gabriels RL, Agnew JA, Miller LJ, et al. Is there a relationship between restricted, repetitive, stereotyped behaviors and interests and abnormal sensory response in children with autism spectrum disorders. *Research in Autism Spectrum Disorders*. 2008; 2:660–670.
11. Gabriels RL, Ivers BJ, Hill DE, Agnew JA, McNeill J. Stability of adaptive behaviors in middle-school children with autism spectrum disorders. *Research in Autism Spectrum Disorders*. 2007; 1(4):291–303.
12. Levy A, Perry A. Outcomes in adolescents and adults with autism: A review of the literature. *Research in Autism Spectrum Disord*. 2011; 5:1271–1282.
13. Lecavalier L, Leone S, Wiltz J. The impact of behaviour problems on caregiver stress in young people with autism spectrum disorders. *J Intellect Disabil Res*. 2006; 50(Pt 3):172–183. [PubMed: 16430729]
14. Cadman T, Eklund H, Howley D, et al. Caregiver burden as people with autism spectrum disorder and attention-deficit/hyperactivity disorder transition into adolescence and adulthood in the United Kingdom. *J Am Child Adolesc Psychiatry*. 2012; 51:879–888.
15. Schieve LA, Blumberg SJ, Rice C, Visser SN, Boyle C. The relationship between autism and parenting stress. *Pediatrics*. 2007; 119(Suppl 1):S114–121. [PubMed: 17272578]
16. Esposito, L.; McCardle, P.; Maholmes, V.; McCune, S.; Griffin, JA. Introduction. In: McCardle, P.; McCune, S.; Griffin, JA.; Esposito, L.; Freund, L., editors. *Animals in our lives: Human-animal interaction in family, community, and therapeutic settings*. Baltimore: Brookes; 2011. p. 1-5.
17. Johnson, RA. Animal-assisted intervention in health care contexts. In: McCardle, P.; McCune, S.; Griffin, JA.; Maholmes, V., editors. *How Animals Affect Us: Examining the Influence of Human-*

Animal Interaction on Child Development and Human Health. Washington D.C: American Psychological Association; 2011. p. 183-192.

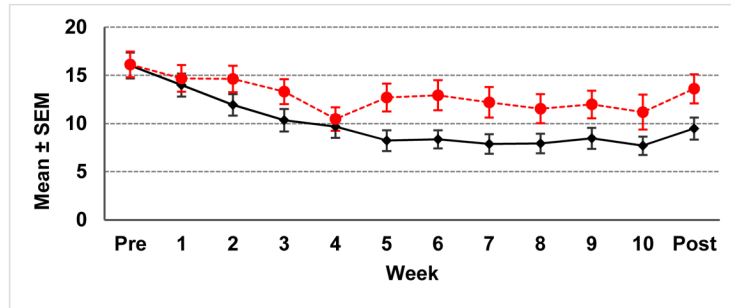
18. Viau R, Aresenault-Lapierre G, Fecteau S, Champagne N, Walker CD, Lupien S. Effect of service dogs on salivary cortisol secretion in autistic children. *Psychoneuroendocrinology*. 2010; 35:1187–1193. [PubMed: 20189722]
19. Pendry P, Smith AN, Roeter SM. Randomized trial effects of equine facilitated learning on adolescent basal cortisol levels. *Human-Animal Interaction Bulletin*. 2014; 2(1):80–95.
20. O’Haire ME, McKenzie SJ, Beck AM, Slaughter V. Social behaviors increase in children with autism in the presence of animals compared to toys. *PLoS ONE*. 2013; 8(2):e57010. [PubMed: 23468902]
21. Bass MM, Duchowny CA, Llabre MM. The effect of therapeutic horseback riding on social functioning in children with autism. *Journal of Autism and Developmental Disorders*. 2009; 39(9): 1261–71. [PubMed: 19350376]
22. Gabriels RL, Agnew JA, Holt KD, et al. Pilot study measuring the effects of therapeutic horseback riding on school-age children and adolescents with autism spectrum disorders. *Research in Autism Spectrum Disorders*. 2012; 6:578–588.
23. Holm MB, Baird JM, Kim YJ, et al. Therapeutic horseback riding outcomes of parent identified goals for children with autism spectrum disorder: An ABA multiple case design examining dosing and generalization to the home and community. *J Autism Dev Disord*. 2013; 44(4):937–47. [PubMed: 24091469]
24. Ward SC, Whalon K, Rusnak K, Wendell K, Paschall N. The association between therapeutic horseback riding and the social communication and sensory reactions of children with autism. *Journal of Autism and Developmental Disorders*. 2013; 43:2190–2198. [PubMed: 23371511]
25. Ajzenman HF, Standeven JW, Shurtleff TL. Effect of hippotherapy on motor control, adaptive behaviors, and participation in children with autism spectrum disorder: A pilot study. *American Journal of Occupational Therapy*. 2013; 67(6):653–63. [PubMed: 24195899]
26. Taylor RR, Keilhofner G, Smith C, Butler S, Cahill SM, Ciukaj MD. Volitional change in children with autism: A single-case design study of the impact of hippotherapy on motivation. *Occupational Therapy in Mental Health*. 2009; 25:192–200.
27. Keino H, Funahashi A, Miwa C, Hosokawa M, Hayashi Y. Psycho-educational horseback riding to facilitate communication ability of children with pervasive developmental disorders. *Journal of Equine Science*. 2009; 20(4):79–88. [PubMed: 24833971]
28. Kern JK, Fletcher CL, Garver CR, Mehta JA, Grannemann BD, Knox KR. Prospective trial of equine-assisted activities in autism spectrum disorder. *Alternative Therapies in Health and Medicine*. 2011; 17(3):14–20. [PubMed: 22164808]
29. Memishevikj H, Hodzhikj S. The effects of equine-assisted therapy in improving the psychosocial functioning of children with autism. *Journal of Special Education and Rehabilitation*. 2010; 11(3–4):57–67.
30. PATH International. [Accessed 24 March 2011] PATH International web site. <http://www.pathintl.org/>. Published 2011
31. Hippotherapy as a treatment strategy. American Hippotherapy Association web site; <http://www.americanhippotherapyassociation.org/hippotherapy/hippotherapy-as-a-treatment-strategy/>. Published 2010 [Accessed December 17, 2011]
32. Grandin, T.; Johnson, C. *Animals in Translation*. New York: Simon and Schuster; 2005.
33. Rutter, M.; Bailey, A.; Lord, C. *Social Communication Questionnaire*. Los Angeles: Western Psychological Services; 2003.
34. Lord, C.; Rutter, M.; DiLavore, PC.; Risi, S. *The autism diagnostic observation schedule ADOS Manual*. Los Angeles: Western Psychological Services; 2003.
35. Lord, C.; Rutter, M.; DiLavore, PC.; Risi, S.; Gotham, K.; Bishop, S. *Autism Diagnostic Observation Schedule. 2*. Los Angeles, CA: Western Psychological Services; 2012. Manual
36. Aman MG, Singh NN, Stewart AW, Field CJ. The aberrant behavior checklist: A behavior rating scale for the assessment of treatment effects. *Am J Ment Defic*. 1985; 89(5):485–491. [PubMed: 3993694]

37. Roid, GH.; Miller, LJ. Leiter International Performance Scale-Revised (Leiter-R). Wood Dale, IL: Stoelting; 1997.
38. Schopler, E.; Mesibov, GB.; Hearsey, K. Structured teaching in the TEACCH system. In: Schopler, E.; Mesibov, GB., editors. Learning and Cognition in Autism. New York: Plenum Press; 1995. p. 243-68.
39. Dunn, LM.; Dunn, DM. Peabody Picture Vocabulary Test, Fourth Edition (PPVT - 4). Austin, TX: Pearson Assessments; 2007.
40. Miller, JF.; Chapman, RS. SALT: A computer program for the Systematic Analysis of Language Transcripts. Madison, WI: University of Wisconsin; 2000.
41. Bruininks, RH.; Bruininks, B. BOT-2: Bruininks-Oseretsky Test of Motor Proficiency. Eagan, MI: Pearson Education; 2005.
42. Ayres, AJ. Sensory Integration and Praxis Tests. Los Angeles: Western Psychological Services; 1989.
43. Sparrow, SS.; Cicchetti, DV.; Balla, DA. Vineland Adaptive Behavior Scales. 2. Circle Pines, MN: American Guidance Service; 2005.
44. Aman MG, Burrow WH, Wolford PL. The Aberrant Behavior Checklist-Community: Factor validity and effect of subject variables for adults in group homes. *Am J Ment Retard.* 1995; 100(3):283-292. [PubMed: 8554775]
45. Aman MG, Richmond G, Stewart AW, Bell JC, Kissel RC. The aberrant behavior checklist: Factor structure and the effect of subject variables in American and New Zealand facilities. *Am J Ment Defic.* 1987; 91(6):570-578. [PubMed: 3591845]
46. Arnold EA, Vitiello B, McDougle C, et al. Parent-Defined Target Symptoms Respond to Risperidone in RUPP Autism Study: Customer Approach to Clinical Trials. *Journal of the American Academy of Child and Adolescent Psychiatry.* 2003; 42(12):1443-50.
47. Constantino, JN. The Social Responsiveness Scale. Los Angeles: Western Psychological Services; 2002.
48. Constantino JN, Davis SA, Todd RD, et al. Validation of a brief quantitative measure of autistic traits: Comparison of the social responsiveness scale with the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders.* 2003; 33(4):427-33. [PubMed: 12959421]
49. SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.
50. Stoner, J. Efficacy of hippotherapy as a treatment strategy for children with autism. In: Engel, B.; MacKinnon, J., editors. Enhancing Human Occupation through Hippotherapy. Bethesda, MD: American Occupational Therapy Association; 2007. p. 103-110.

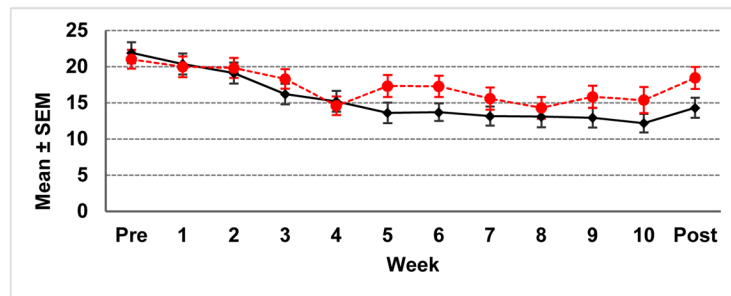


**Figure 1.** Therapeutic horseback riding (THR) and autism spectrum disorder (ASD): participant flowchart. Note: ADOS = Autism Diagnostic Observation Schedule; BA = barn activity; ITT = intent to treat. <sup>a</sup> Eligible for ITT analysis; <sup>b</sup> Eligible for per protocol analysis.

**A) Time course of ABC-C Irritability score**



**B) Time course of ABC-C Hyperactivity score**



---●--- Barn      —◆— THR

**Figure 2.** Time course of Aberrant Behavior Checklist-Community (ABC-C) subscale scores during the course of the 10-week therapeutic horseback riding (THR) intervention. Note: SEM = standard error of the mean.

**Table 1**

## Characteristics of Participants Who Completed the Study

Characteristic	THR	BA Control	Total	<i>p</i> Value <sup>a</sup>
Participants, n	58	58	116	
Age, mean (SD), years	10.5 (3.2)	10.0 (2.7)	10.2 (3.0)	0.34
Gender, M/F (counts)	49/9	52/6	101/15	0.58
IQ, mean (SD)	86.7 (25.5)	86.1 (22.7)	86.4 (24.0)	0.89
Repetitive behavior scale total score, mean (SD)	38.0 (20.9)	38.1 (19.8)	38.0 (20.2)	0.98
Community psychiatric diagnoses, %	48.3	48.3	48.3	1.0
Mood disorder	13.8	20.7	17.2	0.32
Anxiety disorder	27.6	15.5	21.6	0.11
ADHD	29.3	27.6	28.4	0.84
Learning disability	5.2	1.7	3.4	1.0
Current seizure disorder	1.7	3.4	2.6	0.56
Psychotropic medicine	46.6	50.0	48.3	0.71
Distance traveled to riding center, mean (SD)	30.8 (19.1)	27.0 (18.7)	28.9 (18.9)	0.28
Latino/Hispanic	10	11	21	0.77
Race				0.29
American Indian or Alaska native	0	3	3	
Asian	2	2	4	
Black or African American	1	0	1	
Hawaiian/Pacific Islander	0	1	1	
Caucasian	47	45	92	
Multiracial	5	2	7	
Other	1	4	5	
Missing data	2	1	3	

Note: ADHD = attention-deficit/hyperactivity disorder; BA = barn activity; F = female; M = male; THR = therapeutic horseback riding.

<sup>a</sup>Two-tailed *p* value from two sample t-test, chi square test, or Fisher's exact test, as appropriate.

**Table 2**  
 Intent-to-Treat Analysis of Efficacy of Therapeutic Horseback Riding (THR) Intervention Compared to Barn Activity (BA) Control<sup>a</sup>

	THR				BA (Control)				Interaction (efficacy)					
	Baseline		EoT		Change		Baseline		EoT		Change		p	ES
	Mean (SD)	Mean (SD)	Mean (SEM)	Mean (SEM)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SEM)	Mean (SEM)			
ABC														
Irritability <sup>b</sup>	<b>16.0 (9.84)</b>	<b>9.5 (7.98)</b>	<b>-6.3 (1.08)</b>	<b>16.1 (9.80)</b>	<b>13.6 (10.08)</b>	<b>-2.6 (1.1)</b>	<b>16.1 (9.80)</b>	<b>13.6 (10.08)</b>	<b>-2.6 (1.1)</b>	<b>-3.7 (1.53)*</b>	<b>0.02</b>	<b>0.50</b>		
Lethargy/Social Withdrawal	11.5 (8.14)	6.4 (6.70)	-4.8 (0.90)	10.5 (8.65)	7.7 (8.72)	-3.2 (0.9)	10.5 (8.65)	7.7 (8.72)	-3.2 (0.9)	-1.6 (1.28)	0.20	0.26		
Stereotypy	6.3 (5.07)	4.7 (5.06)	-1.4 (0.47)	5.3 (4.50)	3.8 (4.53)	-1.3 (0.5)	5.3 (4.50)	3.8 (4.53)	-1.3 (0.5)	-0.0 (0.66)	0.96	0.01		
Hyperactivity	<b>21.9 (10.75)</b>	<b>14.3 (9.66)</b>	<b>-7.5 (1.25)</b>	<b>21.0 (9.69)</b>	<b>18.4 (10.26)</b>	<b>-2.9 (1.3)</b>	<b>21.0 (9.69)</b>	<b>18.4 (10.26)</b>	<b>-2.9 (1.3)</b>	<b>-4.5 (1.78)*</b>	<b>0.01</b>	<b>0.53</b>		
Inappropriate Speech	4.9 (3.51)	3.1 (3.18)	-1.6 (0.40)	4.3 (2.84)	3.5 (2.90)	-0.9 (0.4)	4.3 (2.84)	3.5 (2.90)	-0.9 (0.4)	-0.7 (0.57)	0.25	0.24		
SRS														
Social Awareness	13.7 (3.16)	12.2 (3.14)	-1.3 (0.37)	13.2 (3.54)	12.4 (3.36)	-1.0 (0.4)	13.2 (3.54)	12.4 (3.36)	-1.0 (0.4)	-0.3 (0.53)	0.54	0.13		
Social Cognition	<b>20.3 (5.63)</b>	<b>17.6 (5.55)</b>	<b>-2.4 (0.66)</b>	<b>19.3 (5.58)</b>	<b>19.1 (5.64)</b>	<b>-0.5 (0.7)</b>	<b>19.3 (5.58)</b>	<b>19.1 (5.64)</b>	<b>-0.5 (0.7)</b>	<b>-1.9 (0.94)*</b>	<b>0.05</b>	<b>0.41</b>		
Social Communication	<b>36.8 (10.04)</b>	<b>30.2 (8.75)</b>	<b>-6.1 (1.11)</b>	<b>33.9 (8.84)</b>	<b>33.6 (11.38)</b>	<b>-1.2 (1.1)</b>	<b>33.9 (8.84)</b>	<b>33.6 (11.38)</b>	<b>-1.2 (1.1)</b>	<b>-4.8 (1.59)**</b>	<b>0.003</b>	<b>0.63</b>		
Autistic Mannerisms	21.2 (6.36)	18.4 (6.04)	-2.5 (0.68)	21.2 (6.30)	19.4 (6.37)	-2.0 (0.7)	21.2 (6.30)	19.4 (6.37)	-2.0 (0.7)	-0.5 (0.97)	0.61	0.10		
Social Motivation	15.8 (5.88)	11.9 (4.97)	-3.6 (0.69)	15.2 (5.09)	13.2 (6.36)	-2.3 (0.7)	15.2 (5.09)	13.2 (6.36)	-2.3 (0.7)	-1.3 (0.99)	0.19	0.27		
BOT-2														
Raw score	48.0 (21.26)	51.4 (22.07)	4.4 (1.06)	50.0 (21.60)	53.6 (21.69)	2.6 (1.1)	50.0 (21.60)	53.6 (21.69)	2.6 (1.1)	1.7 (1.51)	0.26	0.24		
PPVT-4														
Raw score	120.4 (49.24)	126.4 (52.16)	7.8 (1.33)	115.6 (52.38)	125.9 (50.86)	7.1 (1.3)	115.6 (52.38)	125.9 (50.86)	7.1 (1.3)	0.7 (1.89)	0.70	0.08		
SALT														
Number different words used	<b>104.6 (58.45)</b>	<b>116.7 (66.00)</b>	<b>15.7 (5.17)</b>	<b>119.1 (64.55)</b>	<b>118.4 (62.75)</b>	<b>-3.7 (5.3)</b>	<b>119.1 (64.55)</b>	<b>118.4 (62.75)</b>	<b>-3.7 (5.3)</b>	<b>19.4 (7.40)**</b>	<b>0.01</b>	<b>0.54</b>		
Number words used	<b>219.2 (132.19)</b>	<b>253.7 (154.62)</b>	<b>40.5 (14.81)</b>	<b>277.6 (171.53)</b>	<b>270.5 (162.88)</b>	<b>-14.9 (15.1)</b>	<b>277.6 (171.53)</b>	<b>270.5 (162.88)</b>	<b>-14.9 (15.1)</b>	<b>55.3 (21.18)**</b>	<b>0.01</b>	<b>0.54</b>		
Sensory Integration Praxis Tests														
Postural praxis	21.2 (8.72)	23.5 (9.29)	2.6 (0.62)	22.1 (10.33)	23.7 (9.89)	1.2 (0.6)	22.1 (10.33)	23.7 (9.89)	1.2 (0.6)	1.5 (0.88)	0.10	0.35		
Praxis on verbal command	14.6 (8.75)	15.5 (9.18)	1.1 (0.47)	14.8 (8.90)	16.3 (8.25)	1.2 (0.5)	14.8 (8.90)	16.3 (8.25)	1.2 (0.5)	-0.1 (0.67)	0.85	0.04		
VABS-II														
Adaptive total score	70.1 (13.27)	72.4 (14.73)	1.6 (0.94)	70.8 (10.28)	72.4 (11.72)	2.2 (1.0)	70.8 (10.28)	72.4 (11.72)	2.2 (1.0)	-0.7 (1.40)	0.64	0.11		



	THR				BA (Control)				Interaction (efficacy)					
	Baseline		EoT		Change		Baseline		EoT		Change		p	ES
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SEM)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SEM)	Mean (SEM)	Mean (SEM)		
Communication raw score	135.4 (33.75)	140.9 (36.93)	133.3 (34.95)	4.4 (2.00)	139.6 (31.87)	139.6 (31.87)	133.3 (34.95)	139.6 (31.87)	139.6 (31.87)	4.5 (2.2)	-0.1 (2.99)	4.5 (2.2)	0.97	0.01
Daily living raw score	112.3 (38.00)	116.3 (37.72)	112.5 (36.45)	3.7 (2.31)	120.2 (32.58)	120.2 (32.58)	112.5 (36.45)	120.2 (32.58)	120.2 (32.58)	6.1 (2.5)	-2.4 (3.43)	6.1 (2.5)	0.49	0.15
Socialization raw score	95.4 (35.53)	107.0 (37.60)	102.9 (28.55)	9.6 (3.03)	108.8 (30.78)	108.8 (30.78)	102.9 (28.55)	108.8 (30.78)	108.8 (30.78)	6.1 (3.4)	3.5 (4.53)	6.1 (3.4)	0.44	0.17

Note: Significant results are in boldface type. Mean and standard errors of change are from mixed effects model. Effect size (ES) is calculated as  $(2 \times t \text{ value}) / \sqrt{DF}$ . ABS = Aberrant Behavior Checklist; BOT-2 = Bruininks-Oseretsky Test, 2nd Ed.; PPVT-4 = Peabody Picture Vocabulary Test, 4th Ed.; SALT = Systematic Analysis of Language Transcripts; SEM = standard error of the mean; SRS = Social Responsiveness Scale; VABS-II = Vineland Adaptive Behavior Scales-II.

<sup>a</sup> Sample means and standard deviation were reported for baseline and end of treatment (EoT). Mean and standard errors of change are from mixed effects model.

<sup>b</sup> Irritability subscale is deemed as the primary efficacy outcome in this study.

\* p<0.05,

\*\* p<0.01 for testing the time by group interaction in linear mixed effects model analysis. Effect size is calculated as  $(2 \times t \text{ value}) / \sqrt{DF}$ .